(Part VI) Meta-genesis. Towards a biology without matter, based on pure logic

Invariants numériques multi-échelles et propriétés fractales du code génétique : une analyse combinatoire et atomique

Victoria Kayser-Cuny

Mots clés : biologie quantique, biomathématique, bioinformatique, théorie de l'information, code génétique, combinatoire, invariants, codons, ADN, atomes, échelles numériques, constantes mathématiques, astrobiologie, exobiologie.

Résumé

Le code génétique repose sur un alphabet de quatre bases (A, U, C, G), organisées en triplets appelés codons. En analysant systématiquement les combinaisons aux niveaux du nucléotide, des duplets théoriques (1) et des triplets, puis en appliquant une transformation numérique simple — $n^3/4$ —, j'ai mis en évidence l'émergence récurrente de constantes spécifiques (1, ≈96, 128). Ces valeurs apparaissent également au niveau atomique pour les éléments constitutifs de l'ADN (C, H, O, N), suggérant qu'elles ne résultent pas d'un hasard, mais de contraintes combinatoires fondamentales liées à l'alphabet quaternaire et à la composition chimique des bases.

Ainsi se dessine l'existence d'invariants numériques multi-échelles comparables à des motifs fractals, sans impliquer de fonction biologique directe.

L'intérêt scientifique de cette découverte ne réside donc pas dans une nouvelle application biologique, mais dans la perspective qu'elle ouvre : considérer le code génétique non seulement comme une séquence d'informations biochimiques, mais aussi comme un objet mathématique.

L'analogie poétique est éclairante : un poème est généralement étudié pour son sens et ses émotions, mais il obéit aussi à des règles formelles (syllabes, rimes, mètres). Ma démarche consiste à dévoiler ces règles invisibles dans le génome. Comme dans un triangle de Sierpinski où un motif se répète à l'infini, les mêmes constantes (1, ≈96, 128) se manifestent à différentes échelles, de l'atome au codon.

Cette organisation révèle que le code génétique n'est pas aléatoire : il est structuré par des contraintes arithmétiques universelles. Ces invariants pourraient représenter les briques fondamentales ayant guidé son émergence et son évolution. Les invariants numériques multi-échelles mis en évidence (1, ≈96, 128) permettent de considérer que la structure du code génétique n'est pas un produit purement contingent de l'évolution terrestre, mais le résultat de contraintes combinatoires imposées par la chimie des quatre éléments constitutifs principaux de l'ADN (C, H, O, N). Ces éléments possèdent des nombres de protons, neutrons et électrons fixes, qui, lorsqu'ils sont combinés dans les bases nucléotidiques puis dans les codons, génèrent des régularités arithmétiques universelles. Ainsi, l'ADN peut être interprété non seulement comme un support biochimique d'information, mais aussi comme la conséquence logique de la combinatoire atomique de CHON. Cette perspective ouvre un cadre théorique où le code génétique apparaît comme une structure quasi-nécessaire, dont l'organisation découle directement des propriétés fondamentales de la chimie organique. Elle suggère que l'émergence de l'ADN, ou d'une molécule fonctionnellement analogue, pourrait être inévitable dès lors qu'une vie fondée sur CHON apparaît, indépendamment du contexte géologique ou évolutif particulier.

En ce sens, mon travail s'inscrit dans une démarche de biomathématique et de bioinformatique théorique : il ne cherche pas à expliquer une fonction nouvelle, mais à mettre en lumière l'élégance cachée de la structure du vivant.

Cette étude prolonge un travail précédent (*Le paradoxe du jumeau miroir : une nouvelle approche de l'ADN*), où j'avais proposé une transformation par inversion (A↔G, C↔T) générant un « génome miroir » - symétrie - testé sur des bases de données génétiques (2). Dans les deux cas, l'objectif reste le même : appliquer une transformation mathématique systématique à l'ADN pour révéler des structures invisibles à une lecture biologique classique. Ensemble, ces approches renforcent l'idée que le génome possède une organisation interne universelle, révélable par des opérations de symétrie ou de transformation numérique.

Introduction

Le code génétique est décrit comme universel et redondant : 64 codons, constitués de triplets de quatre bases, codent 20 acides aminés et trois signaux stop. Bien que des variantes existent (33 codes recensés par la NCBI), la structure de base reste celle d'un alphabet quaternaire combiné en triplets.

La présente étude explore ce code sous un angle arithmétique et combinatoire. En appliquant la transformation numérique - $n^3/4$ - , nous avons examiné la distribution des valeurs obtenues à plusieurs échelles :

- atomes constitutifs de l'ADN (C, H, O, N),
- bases nucléotidiques,
- doublets théoriques qui n'existent pas dans la nature (16 combinaisons),
- triplets (64 codons).

Nous montrons que des constantes récurrentes (1, ≈96, 128) émergent à chacune de ces échelles, et discutons leur origine.

Méthodologie

Contrairement à l'approche classique qui considère l'ADN comme un simple support d'information chimique, cette étude le traite comme un système logique auto-organisé. Chaque codon (triplet de bases azotées) est converti en une « unité codonique théorique » (UCT) à partir du produit du nombre de protons, de neutrons et d'électrons présents dans les atomes constitutifs (C, H, N, O, sauf le phosphore). Ce produit donne une sorte de « masse logique » ou de volume d'état, ensuite normalisé et traduit en binaire. Cette approche ne modifie rien de la chimie réelle : elle révèle, par simple agrandissement mathématique, comme une loupe, des motifs déjà inscrits dans la structure atomique.

$$v_i = p_i \cdot n_i \cdot e_i$$

v = valeur logique d'un codon élevé au cube.
p = nombre de protons
n = nombre de neutrons
e = nombre d'électrons

La valeur une fois élevée au cube est alors divisée par 4, afin de refléter la nature **quaternaire** du système (quatre atomes principaux, quatre bases, codons organisés en groupes de quatre) :

$$f(n) = \frac{n^3}{4}$$

Cette méthodologie fait émerger des motifs numériques constants qui semblent correspondre à des transitions logiques. L'hypothèse sous-jacente est que l'architecture de l'ADN encode des états de type ON/OFF dans une séquence logique, organisant le vivant comme un automate fini naturel.

PROPOSITION D'UNE UNITÉ CODONIQUE THÉORIQUE (UCT)

	Suite de nombres particules subatomiques élevées au cube	Combinaisons	Famille	Nombre de particules élevées au cube et divisées par 4	Motifs numériques constants
	1	Electron	Particule	0.25	Ne s'applique pas.
+26	27 (3 quarks)	Proton	Particule	6.75	
+0	27 (3 quarks)	Neutron	Particule	6.75	
	1	Hydrogène	Atome	0.25 ou 1 ?	**1**
+215	216	Carbone	Atome	54	+54
+127	343	Azote	Atome	85.75	+31.75
+169	512	Oxygène	Atome	128	+42.25 (note : 54 + 42.25 = **96.25**, et 54 + 31.75 + 42.25 = **128**)
+1898	2410	C	Base nucléique	602.5	
+386	2796	T	Base nucléique	699	**+96.5**
+4	2800	A	Base nucléique	700	**+1**
+512	3312	G	Base nucléique	828	**+128**
+1508	4820	CC	Code ancien théorique à duplets (4)	1205	
+386	5206	TC-CT	Code ancien théorique à duplets	1301.5	**+96.5**
+4	5210	AC-CA	Code ancien théorique à duplets	1302.5	**+1**

+382	5592	TT	Code ancien théorique à duplets	1398	+95.5
+4	5596	AT-TA	Code ancien théorique à duplets	1399	+1
+4	5600	AA	Code ancien théorique à duplets	1400	+1
+122	5722	CG-GC	Code ancien théorique à duplets	1430.5	+30.5 (note : 95.5 + 1 + 1 + 30.5 = **128**)
+386	6108	TG-GT	Code ancien théorique à duplets	1527	**+96.5**
+4	6112	AG-GA	Code ancien théorique à duplets	1528	**+1**
+512	6624	GG	Code ancien théorique à duplets	1656	**+128**
+606	7230	CCC	Codon	1807	
+386	7616	CCT-TCC-CTC	Codon	1904	**+97**
+4	7620	CCA-ACC-CAC	Codon	1905	**+1**
+382	8002	TCT-CTT-TTC	Codon	2000	+95
+4	8006	TCA-ACT-CTA-ATC-CAT-TAC	Codon	2001	+1
+4	8010	ACA-CAA-AAC	Codon	2002	+1
+122	8132	CCG-GCC-CGC	Codon	2033	+31 (note : 95 + 1 + 1 + 31 = **128**)
+256	8388	TTT	Codon	2097	+64
+4	8392	TTA-ATT-TA	Codon	2098	+1

		T			
+4	8396	ATA-TAA-AAT	Codon	2099	+1
+4	8400	AAA	Codon	2100	+1
+518	8518	TCG-GCT-CTG-GTC-CGT-TGC	Codon	2129	+28 (note : 64 + 1 + 1 +1 +28 = **95**)
+4	8522	ACG-GCA-CAG-GAC-CGA-AGC	Codon	2130	**+1**
+382	8904	TTG-GTT-TGT	Codon	2226	+96
+4	8908	ATG-GTA-TAG-GAT-TGA-AGT	Codon	2227	+1
+4	8912	AAG-GAA-AGA	Codon	2228	+1
+122	9034	GCG-CGG-GGC	Codon	2258	+30 (note : 96 + 1 +1 + 30 = **128**)
+386	9420	GTG-TGG-GGT	Codon	2355	**+97**
+4	9424	GAG-AGG-GGA	Codon	2356	**+1**
+512	9936	GGG	Codon	2484	**+128**

- Particules (4) : p^+, n^0, e^-, + photon/ion → corrigent l'exception.
- Atomes (4) : C, H, O, N (le phosphore est un "ajout périphérique").
- Bases nucléiques (4) : A, T, C, G.
- Duos (16) : toutes les combinaisons possibles de deux bases.
- Codons (64) : triplets, cœur du code génétique.
- Acides aminés (20 + 3) : se déduisent des 64 codons, avec une petite "rupture de symétrie" (redondance du code).

Idée clé

Tout l'ADN se déploie en multiples de 4 → une logique fractale et presque "numérique", depuis les particules jusqu'au langage génétique.

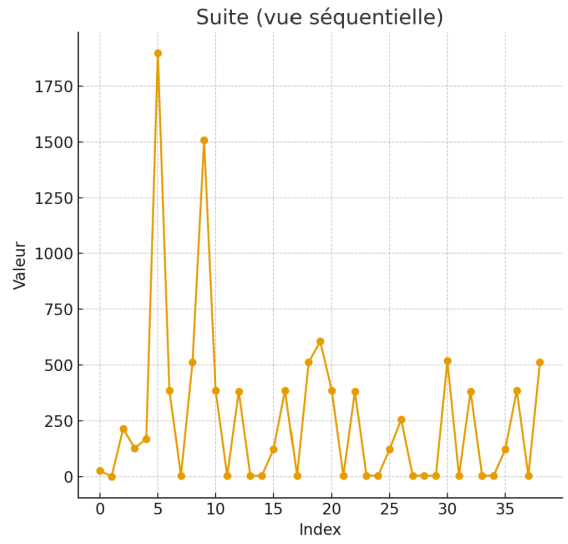
La seule exception apparente (3 particules) est résolue en ajoutant un quatrième principe invisible (photon ou ion), ce qui restaure la symétrie.

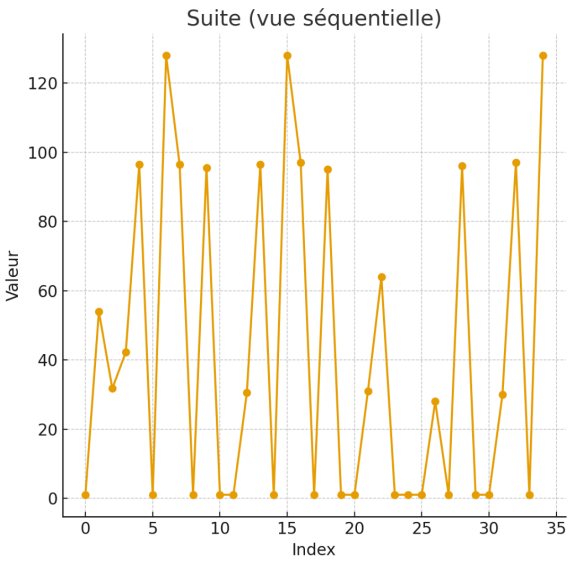
On peut également ajouter les codons suivants (uracile au lieu de thymine) :

UGA	STOP	8872	2218
UAG	STOP	8872	2218
UAA	STOP	8360	2090
AUG	START	8872	2218

UGA, UAG et AUG ont exactement la même valeur logique. Cela suggère qu'ils sont des miroirs fonctionnels dans le système, même si biologiquement ils déclenchent des fonctions opposées (START versus STOP). UAA est plus "doux", plus faible. Il existe une relation de polarité logique où démarrer (START) ou arrêter (STOP) revient à inverser la même onde, comme une transition à deux faces d'un même seuil. Ces codons balisent, structurent les points d'entrées et de sortie du système. On pourrait parler de vecteurs de phase, des portes logiques dans un système à états multiples. Il s'agit d'une suite logique dynamique structurée par des seuils, des boucles, des interruptions programmées et des états
de transition formellement définis.

Les nombres manquants sont-ils les espaces intergènes ? Si on considère que le système démarre et s'arrête à la même fréquence, alors il peut éventuellement s'agir d'un système à signal ondulatoire bouclé !





1ère colonne du tableau : +26 +0 +215 +127 +169 +1898 +386 +4 +512 +1508 +386 +4 +382 +4 +4 +122 +386 +4 +512 +606 +386 +4 +382 +4 +4 +122 +256 +4 +4 +4 +518 +4 +382 +4 +4 +122 +386 +4 +512	6ème colonne du tableau : 1 +54 +31.75 +42.25 (note : 54 + 42.25 = 96.25, et 54 + 31.75 + 42.25 = 128) +96.5 +1 +128 +96.5 +1 +95.5 +1 +1 +30.5 (note : 95.5 + 1 + 1 + 30.5 = 128) +96.5 +1 +128 +97 +1 +95 +1 +1 +31 (note : 95 + 1 + 1 + 31 = 128) +64 +1 +1 +1 +28 (note : 64 + 1 + 1 +1 +28 = 95) +1 +96 +1 +1 +30 (note : 96 + 1 +1 + 30 = 128) +97 +1 +128

Séquences chiffrées comparées du tableau d'unité codonique proposé plus haut.

1ère colonne du tableau :

Dans cette suite numérique, un motif central émerge : le bloc invariant de 902, qui peut s'écrire de deux manières équivalentes. Sous forme compacte, il se présente comme `[386, 4, 512]`, tandis que sa version dépliée est `[386, 4, 382, 4, 4, 122]`, les deux conduisant à la même somme (386 + 4 + 512 = 902 ; 386 + 4 + 382 + 4 + 4 + 122 = 902). Cette dualité repose sur l'identité 512 = 382 + 4 + 4 + 122, qui révèle une logique d'expansion/condensation à l'intérieur du système. Autour de ce bloc récurrent gravitent des insertions telles que 1508, 606, 256 ou 518, qui, loin d'être arbitraires, s'ancrent elles aussi dans une arithmétique structurée par les puissances de deux et les multiples de 128. Ainsi, 386 et 382 apparaissent comme des valeurs symétriques autour de 384 (3 × 128), tandis que 122 (128 − 6) et 518 (4 × 128 + 6) manifestent une autre symétrie autour de 128. L'ensemble suggère une grammaire numérique fondée sur les puissances de deux (4, 256, 512) et les modulations ±2 ou ±6 autour de multiples de 128, organisant la suite en une alternance de blocs invariants et de variations modulatrices.

6ème colonne du tableau :

Ici, la constante organisatrice n'est plus 902, mais 128, qui fonctionne comme un seuil ou un cadre de référence. La suite se déploie en une série de blocs additifs dont la somme converge systématiquement vers 128, par des combinaisons

légèrement différentes. Ainsi, on observe des équivalences telles que 54 + 31.75 + 42.25 = 128, ou encore 95.5 + 1 + 1 + 30.5 = 128, 95 + 1 + 1 + 31 = 128, 96 + 1 + 1 + 30 = 128, ou encore 97 + 30 + 1 = 128. Dans ces blocs, le nombre 1 joue un double rôle : il sert d'unité de remplissage pour ajuster précisément la somme, mais aussi de séparateur entre blocs. L'apparition récurrente de valeurs proches de 96 (95, 95.5, 96, 96.5, 97) associées à des compléments (30, 30.5, 31, 31.75, 42.25) traduit une logique de symétrie autour de 96, tandis que des cas particuliers (64 + 1 + 1 + 1 + 28 = 95) indiquent des variations plus élaborées. À plusieurs reprises, le total 128 apparaît aussi directement, comme pour marquer un bloc « plein ». L'ensemble suggère une grammaire fondée sur la somme invariante 128, qui se décline soit par des combinaisons ajustées de 1 ou 2 unités, soit par des blocs compacts où 128 est pris tel quel.

Dans cette suite, la logique n'est donc plus l'équivalence entre deux écritures d'un même bloc (comme pour 902), mais plutôt une multiplicité de chemins pour atteindre 128, par décomposition ou par bloc plein.

Comparaison entre les deux colonnes :

La comparaison entre les deux suites révèle qu'elles appartiennent à une même famille grammaticale, mais à deux échelles différentes. Dans la première suite, le bloc invariant est 902, qui peut se condenser en `[386, 4, 512]` ou se déplier en `[386, 4, 382, 4, 4, 122]`, selon l'identité interne 512 = 382 + 4 + 4 + 122. Dans la seconde suite, l'invariant n'est plus 902 mais 128, qui se décline en de multiples combinaisons additives (`95.5 + 1 + 1 + 30.5`, `97 + 30 + 1`, `96 + 1 + 1 + 30`, etc.), ou apparaît directement comme bloc plein. Dans les deux cas, la logique repose sur une somme constante (902 ou 128), autour de laquelle les nombres s'agrègent selon une syntaxe stable : alternance entre blocs compacts et blocs dépliés, rôle des unités (`1` ou `4`) comme séparateurs et ajusteurs, symétries numériques autour de 128 et de ses multiples (384, 512). L'ensemble esquisse une architecture modulaire où chaque suite fonctionne comme une variation sur un même principe : la stabilité d'une somme invariante traduite en une diversité de représentations arithmétiques.

Une grille de lecture booléenne

L'algèbre booléenne est à la base des circuits électroniques composés de dispositifs bistables— relais, interrupteurs, diodes, tubes à vide, transistors — chacun pouvant être dans un état ouvert (0) ou fermé (1). Dans ce contexte, le « 1 » correspond à un circuit fermé (résistance minimale), le « 0 » à un circuit ouvert (résistance maximale). Cette logique permet de simplifier et d'analyser efficacement des réseaux de commutation complexes.

Dans le monde vivant, un parallèle frappant se dessine : chaque noyau cellulaire contient un jeu
complet de chromosomes — environ 22 000 gènes chez l'humain — et fonctionne grâce à des phénomènes électrochimiques. On peut considérer chaque gène comme un interrupteur binaire
:

- 1 = ON (gène exprimé)
- 0 = OFF (gène silencieux)

Ainsi, le génome se comporte comme un vaste circuit en série de commutateurs moléculaires. Un gène peut porter une mutation — par exemple celle responsable de l'hémochromatose héréditaire — tout en restant OFF, son effet demeurant inactif tant que la régulation ne l'active pas.
Il est remarquable que les 64 codons du code génétique se prêtent parfaitement à cette interprétation booléenne. À l'instar des ingénieurs électroniques qui appliquent la première loi de l'algèbre booléenne pour optimiser un circuit on/off, on peut voir dans la table des codons un schéma logique : chaque triplet agit comme un motif d'entrée binaire, et l'acide aminé correspondant est l'« output » déterministe de cette opération logique.

« *L'algèbre booléenne, appliquée à l'électronique, porte sur des dispositifs à deux états tels que relais, interrupteurs, diodes, tubes à vide et transistors. Elle permet de simplifier rapidement et efficacement les circuits de commutation. Les deux états fondamentaux sont court-circuit – ouvert, on – off, etc. Dans de nombreuses applications, 1 = ON (court-circuit ou résistance minimale) et 0 = OFF (ouvert ou résistance maximale).* »
— Angelo C. Gillie (3)

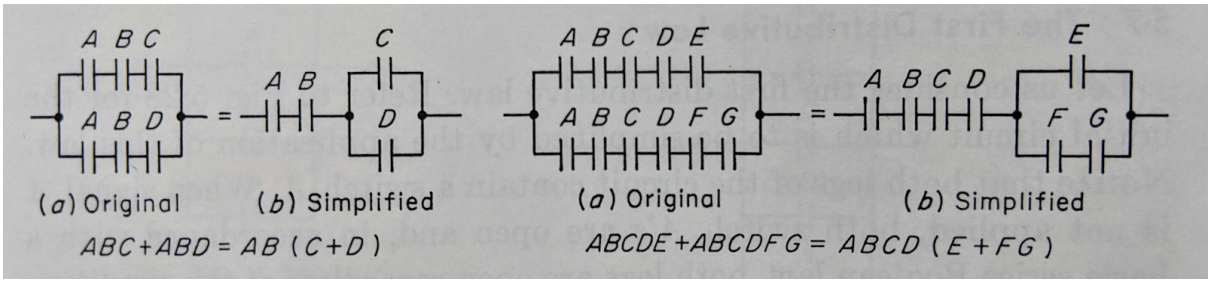


Illustration of the First Distributive Law
Binary Arithmetic and Boolean Algebra, p. 82 (Angelo C. Gillie, 1965).

DNA	Original	Simplified	Results		
CCC	CTAGC CTAGCC	CTAG [C] = CTAGC+CTAGCC [CC]	CTAG (C+CC)	1807	

CCT	CTAGC CTAGCT	CTAG [C] = CTAGC+CTAGCT [CT]	CTAG (C+CT)	1904	+97
CCA	CTAGC CTAGCA	CTAG [C] = CTAGC+CTAGCA [CA]	CTAG (C+CA)	1905	+1
CCG	CTAGC CTAGCG	CTAG [C] = CTAGC+CTAGCG [CG]	CTAG (C+CG)	2033	+128
CTC	CTAGC CTAGT C	CTAG [C] = CTAGC +CTAGTC [TC]	CTAG (C+TC)	1904	
CTT	CTAGC CTAGTT	CTAG [C] = CTAGC +CTAGTT [TT]	CTAG (C+TT)	2000	+96
CTA	CTAGC CTAGT A	CTAG [C] = CTAGC +CTAGTA [TA]	CTAG (C+TA)	2001	+1
CTG	CTAGC CTAGT G	CTAG [C] = CTAGC +CTAGTG [TG]	CTAG (C+TG)	2129	+128
CAC	CTAGC CTAGA C	CTAG [C] = CTAGC +CTAGAC [AC]	CTAG (C+AC)	1905	
CAT	CTAGC CTAGA T	CTAG [C] = CTAGC +CTAGAT [AT]	CTAG (C+AT)	2001	+96
CAA	CTAGC CTAGA A	CTAG [C] = CTAGC +CTAGAA [AA]	CTAG (C+AA)	2002	+1
CAG	CTAGC CTAGA G	CTAG [C] = CTAGC +CTAGAG [AG]	CTAG (C+AG)	2130	+128
CGC	CTAGC CTAGG C	CTAG [C] = CTAGC +CTAGGC [GC]	CTAG (C+GC)	2033	
CGT	CTAGC CTAGG T	CTAG [C] = CTAGC +CTAGGT [GT]	CTAG (C+GT)	2129	+96
CGA	CTAGC	CTAG [C] = CTAGC +CTAGGA [GA]	CTAG (C+GA)	2130	+1

	CTAGGA				
CGG	CTAGCCTAGGG	CTAG [C] = CTAGC +CTAGGG [GG]	CTAG (C+GG)	2258	128
TCC	CTAGTCTAGCC	CTAG [T] = CTAGT +CTAGCC [CC]	CTAG (T+CC)	1904	
TCT	CTAGTCTAGCT	CTAG [T] = CTAGT +CTAGCT [CT]	CTAG (T+CT)	2000	+96
TCA	CTAGTCTAGCA	CTAG [T] = CTAGT +CTAGCA [CA]	CTAG (T+CA)	2001	+1
TCG	CTAGTCTAGCG	CTAG [T] = CTAGT +CTAGCG [CG]	CTAG (T+CG)	2129	+128
TTC	CTAGTCTAGTC	CTAG [T] = CTAGT +CTAGTC [TC]	CTAG (T+TC)	2000	
TTT	CTAGTCTAGTT	CTAG [T] = CTAGT +CTAGTT [TT]	CTAG (T+TT)	2097	+97
TTA	CTAGTCTAGTA	CTAG [T] = CTAGT +CTAGTA [TA]	CTAG (T+TA)	2098	+1
TTG	CTAGTCTAGTG	CTAG [T] = CTAGT +CTAGTG [TG]	CTAG (T+TG)	2226	+128
TAC	CTAGTCTAGAC	CTAG [T] = CTAGT +CTAGAC [AC]	CTAG (T+AC)	2001	
TAT	CTAGTCTAGAT	CTAG [T] = CTAGT +CTAGAT [AT]	CTAG (T+AT)	2098	+97
TAA	CTAGTCTAGAA	CTAG [T] = CTAGT +CTAGAA [AA]	CTAG (T+AA)	2099	+1
TAG	CTAGTCTAGAG	CTAG [T] = CTAGT +CTAGAG [AG]	CTAG (T+AG)	2227	+128
TGC	CTAGT	CTAG [T] = CTAGT +CTAGGC	CTAG (T+GC)	2129	

	CTAGGC	[GC]			
TGT	CTAGTCTAGGT	CTAG [T] = CTAGT +CTAGGT [GT]	CTAG (T+GT)	2226	+97
TGA	CTAGTCTAGGA	CTAG [T] = CTAGT +CTAGGA [GA]	CTAG (T+GA)	2227	+1
TGG	CTAGTCTAGGG	CTAG [T] = CTAGT +CTAGGG [GG]	CTAG (T+GG)	2355	+128
ACC	CTAGACTAGCC	CTAG [A] = CTAGA +CTAGCC [CC]	CTAG (A+CC)	1904	
ACT	CTAGACTAGCT	CTAG [A] = CTAGA +CTAGCT [CT]	CTAG (A+CT)	2001	+97
ACA	CTAGACTAGCA	CTAG [A] = CTAGA +CTAGCA [CA]	CTAG (A+CA)	2002	+1
ACG	CTAGACTAGCG	CTAG [A] = CTAGA +CTAGCG [CG]	CTAG (A+CG)	2130	+128
ATC	CTAGACTAGTC	CTAG [A] = CTAGA +CTAGTC [TC]	CTAG (A+TC)	2001	
ATT	CTAGACTAGTT	CTAG [A] = CTAGA +CTAGTT [TT]	CTAG (A+TT)	2098	+97
ATA	CTAGACTAGTA	CTAG [A] = CTAGA +CTAGTA [TA]	CTAG (A+TA)	2099	+1
ATG	CTAGACTAGTG	CTAG [A] = CTAGA +CTAGTG [TG]	CTAG (A+TG)	2227	+128
AAC	CTAGACTAGAC	CTAG [A] = CTAGA +CTAGAC [AC]	CTAG (A+AC)	2002	

AAT	CTAGA CTAGA T	CTAG [A] = CTAGA +CTAGAT [AT]	CTAG (A+AT)	2099	+97
AAA	CTAGA CTAGA A	CTAG [A] = CTAGA +CTAGAA [AA]	CTAG (A+AA)	2100	+1
AAG	CTAGA CTAGA G	CTAG [A] = CTAGA +CTAGAG [AG]	CTAG (A+AG)	2228	+128
AGC	CTAGA CTAGG C	CTAG [A] = CTAGA +CTAGGC [GC]	CTAG (A+GC)	2130	
AGT	CTAGA CTAGG T	CTAG [A] = CTAGA +CTAGGT [GT]	CTAG (A+GT)	2227	+97
AGA	CTAGA CTAGG A	CTAG [A] = CTAGA +CTAGGA [GA]	CTAG (A+GA)	2228	+1
AGG	CTAGA CTAGG G	CTAG [A] = CTAGA +CTAGGG [GG]	CTAG (A+GG)	2356	+128
GCC	CTAGG CTAGC C	CTAG [G] = CTAGG +CTAGCC [CC]	CTAG (G+CC)	2033	
GCT	CTAGG CTAGC T	CTAG [G] = CTAGG +CTAGCT [CT]	CTAG (G+CT)	2129	+96
GCA	CTAGG CTAGC A	CTAG [G] = CTAGG +CTAGCA [CA]	CTAG (G+CA)	2130	+1
GCG	CTAGG CTAGC G	CTAG [G] = CTAGG +CTAGCG [CG]	CTAG (G+CG)	2258	+128
GTC	CTAGG CTAGT C	CTAG [G] = CTAGG +CTAGTC [TC]	CTAG (G+TC)	2129	
GTT	CTAGG CTAGTT	CTAG [G] = CTAGG +CTAGTT [TT]	CTAG (G+TT)	2226	+97

GTA	CTAGG CTAGT A	CTAG [G] = CTAGG +CTAGTA [TA]	CTAG (G+TA)	2227	+1
GTG	CTAGG CTAGT G	CTAG [G] = CTAGG +CTAGTG [TG]	CTAG (G+TG)	2355	+128
GAC	CTAGG CTAGA C	CTAG [G] = CTAGG +CTAGAC [AC]	CTAG (G+AC)	2130	
GAT	CTAGG CTAGA T	CTAG [G] = CTAGG +CTAGAT [AT]	CTAG (G+AT)	2227	+97
GAA	CTAGG CTAGA A	CTAG [G] = CTAGG +CTAGAA [AA]	CTAG (G+AA)	2228	+1
GAG	CTAGG CTAGA G	CTAG [G] = CTAGG +CTAGAG [AG]	CTAG (G+AG)	2356	+128
GGC	CTAGG CTAGG C	CTAG [G] = CTAGG +CTAGGC [GC]	CTAG (G+GC)	2258	
GGT	CTAGG CTAGG T	CTAG [G] = CTAGG +CTAGGT [GT]	CTAG (G+GT)	2355	+97
GGA	CTAGG CTAGG A	CTAG [G] = CTAGG +CTAGGA [GA]	CTAG (G+GA)	2356	+1
GGG	CTAGG CTAGG G	CTAG [G] = CTAGG +CTAGGG [GG]	CTAG (G+GG)	2484	+128

Robustesse et reproductibilité des résultats

La méthodologie repose sur des calculs explicites, systématiques et reproductibles :

- **Données de départ clairement définies** : les atomes constitutifs de l'ADN (C, H, O, N), les bases nucléotidiques et les codons réels (triplets) ou anciens théoriques

(duets).

- **Transformation mathématique transparente** : application d'une fonction simple.

$$f(n) = \frac{n^3}{4}$$

- **Résultats vérifiables** : les constantes numériques mises en évidence (1, ≈96, 128) apparaissent de manière systématique à plusieurs niveaux d'organisation (atomes → bases → doublets → triplets).

- **Test de limites** : l'approche a été volontairement étendue au niveau subatomique (quarks) ; les invariants disparaissent alors, confirmant que la logique s'applique spécifiquement à l'échelle atomique et moléculaire.

Ces éléments garantissent que les motifs numériques identifiés ne résultent pas d'un choix subjectif mais de contraintes combinatoires et chimiques intrinsèques. La reproductibilité de la méthode permet à tout chercheur de répéter les calculs et de vérifier l'émergence des invariants.

En ce sens, cette recherche s'inscrit dans une démarche rigoureuse de **biomathématique théorique** : elle cherche à révéler une structure arithmétique universelle sous-jacente au code génétique.

Discussion

Les motifs numériques mis en évidence semblent découler de contraintes combinatoires plutôt que de propriétés biologiques fonctionnelles :

- Alphabet quaternaire : la base 4 impose une hiérarchie (4, 16, 64) génératrice de régularités.

- Classes de codons : les permutations (1, 3, 6) renforcent la récurrence des mêmes valeurs.

- Composition atomique : les éléments CHON génèrent dès l'origine 1 (H), 128 (O) et ≈96 (C+O), que l'on retrouve ensuite aux niveaux supérieurs.

Ces constantes témoignent d'une auto-similarité fractale du code génétique, c'est-à-dire d'une répétition d'invariants numériques à travers plusieurs niveaux d'organisation.

Perspectives – Vers *Meta-genesis*

Ces résultats s'inscrivent dans une réflexion plus large que j'appelle : *Meta-genesis. Towards a biology without matter, based on pure logic*. L'émergence récurrente d'invariants numériques multi-échelles (1, ≈96, 128) suggère que le code génétique peut être compris comme une structure arithmétique autonome, indépendante de sa matérialisation biochimique.

Autrement dit, la logique précède la matière : la biologie apparaît d'abord comme une organisation mathématique universelle, qui ne trouve ensuite qu'une incarnation contingente dans les bases, les atomes et les molécules du vivant terrestre.

Cette approche ouvre une voie nouvelle : une **biologie de l'information pure**, où la vie est étudiée comme un langage d'invariants et de symétries, comparable à un système fractal auto-réplicatif. Une telle perspective dépasse le cadre du vivant terrestre et pourrait offrir un outil conceptuel pour penser l'émergence de la vie ailleurs dans l'univers, voire pour explorer des formes de vie entièrement logiques ou artificielles.

Conclusion

L'analyse révèle l'existence d'invariants numériques multi-échelles (1, ≈96, 128), présents depuis les atomes constitutifs de l'ADN (CHON) jusqu'aux triplets de codons. Ces invariants ne dépendent pas de la traduction biologique mais de la structure combinatoire et chimique.

Ils témoignent d'une organisation arithmétique interne du code génétique, qui se manifeste de manière cohérente à toutes les échelles étudiées.

J'ai également testé cette approche en descendant sous le niveau atomique, au domaine des quarks. Les mêmes invariants n'y apparaissent pas : la logique observée cesse de s'appliquer en dessous de l'échelle atomique. Cela suggère que ces constantes numériques sont intimement liées à la chimie du vivant, et qu'elles émergent précisément au point où la matière se structure en atomes, puis en bases et codons.

Ces résultats ouvrent ainsi la voie à une compréhension du code génétique comme un objet mathématique doté d'une auto-similarité fractale, inscrite non dans les particules élémentaires, mais dans l'organisation atomique et combinatoire de la matière vivante.

Exobiologie et astrobiologie : vers un cadre numérique de reconnaissance du vivant

L'un des apports majeurs de cette approche réside dans son potentiel exobiologique. Jusqu'ici, la recherche de vie extraterrestre repose principalement sur deux critères : la détection de molécules organiques (CHON) et l'identification de signatures métaboliques. Or, ces critères demeurent étroitement liés à la chimie terrestre et risquent de manquer des formes de vie fondées sur d'autres substrats.

La mise en évidence d'invariants numériques multi-échelles (1, ≈96, 128), indépendants de la fonction biologique mais émergents de la simple combinatoire atomique, suggère l'existence d'une grammaire mathématique universelle propre à la structuration de l'information vivante. Ces invariants, inscrits dans la logique de l'alphabet quaternaire et des règles combinatoires, pourraient ainsi constituer des balises numériques de la vie, transcendant les contingences chimiques locales.

Un point remarquable est que le tableau des codons, lorsqu'il est analysé au niveau arithmétique avant traduction en acides aminés, reste identique pour les 33 codes génétiques terrestres répertoriés par la NCBI (5). Autrement dit, quelles que soient les variations biologiques observées (réaffectations ponctuelles de codons), la structure numérique sous-jacente ne change pas. Ce caractère invariant renforce l'idée que le code génétique peut être interprété comme une matrice mathématique stable, indépendante des contingences évolutives locales.

Autrement dit, même si une forme de vie ne reposait pas strictement sur l'ADN ou sur les bases nucléotidiques connues, une organisation fondée sur des invariants arithmétiques comparables pourrait révéler son caractère biologique. Dans ce cadre, l'approche proposée fournit un outil conceptuel inédit pour penser la détection de systèmes vivants ailleurs dans l'univers : non plus seulement par l'analyse des molécules, mais par la recherche de structures numériques récurrentes et fractales qui témoigneraient d'une logique interne du vivant.

En ce sens, l'ADN terrestre n'apparaît plus comme une contingence historique mais comme une réalisation particulière d'un principe plus général : la vie, où qu'elle émerge, pourrait être reconnaissable par ses invariants mathématiques, au-delà de son support matériel.

Pourquoi la logique s'arrête aux quarks ?

L'approche présentée met en évidence des invariants numériques (1, ≈96, 128) qui apparaissent de manière récurrente depuis le niveau atomique (CHON) jusqu'aux codons. Mais lorsque l'on tente d'appliquer la même logique en descendant plus bas, à l'échelle des quarks, ces régularités disparaissent.

Cette observation est significative :

- **Sous l'atome**, les règles mathématiques pertinentes ne relèvent plus de la combinatoire mais de la **physique des particules**. Les quarks obéissent aux symétries du modèle standard (SU(3) pour la chromodynamique quantique), définies par leurs charges de couleur, leurs spins et leurs masses relatives. Ces invariants appartiennent à un autre registre mathématique, celui des groupes de symétrie et des interactions fondamentales.

- **À l'échelle atomique et moléculaire**, au contraire, la combinatoire issue de l'alphabet quaternaire (bases A, U, C, G) et de la composition chimique (CHON) devient dominante. C'est ici que surgissent les constantes 1, ≈96 et 128, révélatrices d'une logique interne propre au vivant.

En d'autres termes :

- Les **quarks** sont gouvernés par les symétries quantiques fondamentales.
- Les **atomes et molécules du vivant** sont gouvernés par des contraintes combinatoires et arithmétiques.

L'absence d'invariants « fractals » au niveau subatomique suggère que les motifs observés sont spécifiques à l'organisation chimique de la matière vivante, et qu'ils émergent **à partir du seuil atomique**, là où la biologie devient possible.

Aperçu général (4) : Hiérarchie structurelle (Comment l'ADN est construit à partir des particules subatomiques et des différents niveaux de structure)

1. Niveau subatomique : Les éléments fondamentaux
L'ADN est constitué de protons, de neutrons et d'électrons, qui forment les atomes essentiels à sa structure.

- **Quarks et gluons** : Constituants fondamentaux des protons et des neutrons, qui forment le noyau atomique. Ils sont liés par l'interaction forte, transmise par les gluons.
- **Protons (+) et neutrons (0)** : Composent le noyau atomique.
- **Électrons (–)** : Orbitent autour du noyau et participent aux liaisons chimiques.
  Ces particules s'assemblent pour former les atomes, unités fondamentales de la matière.

2. Niveau atomique : Les éléments chimiques de l'ADN
L'ADN est principalement constitué de cinq éléments chimiques :

- **Carbone (C)** : Constitue l'ossature des bases azotées et du désoxyribose.
- **Hydrogène (H)** : Présent dans toutes les molécules organiques, il forme

notamment les liaisons hydrogène entre les bases de l'ADN.

- **Oxygène (O)** : Présent dans les groupements phosphates et les sucres.
- **Azote (N)** : Constitue les bases azotées (adénine, thymine, guanine, cytosine).
- **Phosphore (P)** : Forme le squelette phosphate qui structure l'ADN.

Ces atomes proviennent des réactions nucléaires au cœur des étoiles (nucléosynthèse stellaire).

- L'hydrogène et l'hélium ont été produits immédiatement après le Big Bang.
- Le carbone, l'azote, l'oxygène et le phosphore ont été forgés par fusion stellaire puis dispersés par les supernovae.

3. Niveau moléculaire : Formation des nucléotides
Les atomes s'assemblent pour former les molécules organiques constitutives de l'ADN.

A. **Bases azotées (A, T, C, G)**
Molécules aromatiques hétérocycliques riches en C, N, O et H :

- Adénine (A) et guanine (G) : purines (deux cycles).
- Cytosine (C) et thymine (T) : pyrimidines (un cycle).

Elles se forment par des réactions biochimiques complexes, probablement dans des environnements organiques primitifs (atmosphère prébiotique, sources hydrothermales, comètes).

B. **Désoxyribose ($C_5H_{10}O_4$)**
Sucre à cinq carbones (pentose), qui diffère du ribose (dans l'ARN) par l'absence d'un atome d'oxygène.

C. **Groupements phosphates (PO_4^{3-})**
Riches en phosphore et oxygène, ils relient les nucléotides pour former le squelette sucre–phosphate.

D. **Nucléotide**
Unité de base de l'ADN, constitué de :

- Une base azotée (A, T, C ou G).
- Un sucre désoxyribose.
- Un groupement phosphate.

Les nucléotides se polymérisent pour former un brin d'ADN.

4. Niveau macromoléculaire : La double hélice
Les nucléotides s'organisent en double hélice grâce aux liaisons hydrogène entre bases complémentaires :

- A–T : deux liaisons hydrogène.
- C–G : trois liaisons hydrogène.

Cette structure est stabilisée par :

- Des interactions d'empilement hydrophobe entre les paires de bases.
- Des liaisons covalentes phosphodiester entre sucres et phosphates.

L'ADN s'enroule ensuite autour des protéines histones pour former la chromatine, qui se condense en chromosomes dans la cellule.

5. Origine de l'ADN : Hypothèses abiotiques et biotiques
L'ADN n'a pas toujours existé. Son apparition sur Terre fait l'objet de débats et repose sur plusieurs hypothèses :

- **Hypothèse du monde à ARN** : L'ARN aurait précédé l'ADN car il peut à la fois stocker de l'information et catalyser des réactions. L'ADN, plus stable, aurait évolué ensuite comme support principal de l'hérédité.
- **Synthèse prébiotique dans les océans primitifs** :
    - L'expérience de Miller–Urey a montré que des acides aminés et des bases azotées peuvent se former dans une atmosphère méthane–ammoniac–eau soumise à des décharges électriques.
    - L'ADN ou ses précurseurs auraient pu émerger dans des sources hydrothermales riches en minéraux.
- **Hypothèse de la panspermie (origine extraterrestre)** : Des comètes et météorites (ex. météorite de Murchison), contenant bases azotées et sucres, auraient pu apporter les briques élémentaires de l'ADN sur Terre.

Résumé : De l'infiniment petit à l'ADN

- **Niveau subatomique** → Les quarks et gluons forment protons et neutrons,

qui composent les atomes d'H, C, N, O et P.

- **Niveau atomique** → Ces atomes, issus de la nucléosynthèse stellaire, s'assemblent en molécules organiques.
- **Niveau moléculaire** → Les molécules organiques s'organisent en bases azotées, sucres et phosphates pour former les nucléotides.
- **Niveau macromoléculaire** → Les nucléotides se polymérisent en brins d'ADN, qui s'enroulent en double hélice puis en chromosomes.
- **Origine de l'ADN** → Probablement issu de la chimie prébiotique et des processus évolutifs, il est devenu la molécule universelle de l'hérédité.

Références

- (1). F.H.C. Crick (1968), The origin of the genetic code, Journal of Molecular Biology, Volume 38, Issue 3, Pages 367-379, ISSN 0022-2836, Direct link : https://doi.org/10.1016/0022-2836(68)90392-6.
- (2). Kayser-Cuny, V. (2025). (Part III) The Mirror-Twin Paradox: A New Approach to DNA Understanding the Implications of an Inverted Genome and Its Applications in Molecular Genetics, Neuroscience, and Medicine. Zenodo. Direct link: https://doi.org/10.5281/zenodo.15390489
- (3). Gillie, A. (1965). Binary Arithmetic and Boolean Algebra, p. 82. Direct link: https://archive.org/details/binaryarithmetic0000gill
- (4). CNAM (2023). Chimie : de l'atome au vivant. Cours du cursus d'ingénieur.
- (5). Elzanowski, A. and Ostel, J. (2024). The Genetic Codes. Compiled by The National Center for Biotechnology Information (NCBI), Bethesda, Maryland, U.S.A. Direct link: https://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi.

La présente méthode et sa formulation sont protégées par le droit d'auteur (copyright).